

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

CLIPPEDIMAGE= JP409083856A

PAT-NO: JP409083856A

DOCUMENT-IDENTIFIER: JP 09083856 A

TITLE: INTELLIGENT CAMERA EQUIPMENT

PUBN-DATE: March 28, 1997

INVENTOR-INFORMATION:

NAME

AOYANAGI, SHIGEKI  
HAGITA, NORIHIRO  
ONOUCHI, RIKIO

ASSIGNEE-INFORMATION:

NAME	COUNTRY
NIPPON TELEGR & TELEPH CORP <NTT>	N/A

APPL-NO: JP07229910

APPL-DATE: September 7, 1995

INT-CL (IPC): H04N005/232;G06T001/00 ;G08B021/00 ;G08B025/00 ;G10L003/00  
;H04N007/15 ;H04N007/18

ABSTRACT:

PROBLEM TO BE SOLVED: To eliminate the need for an exclusive operator by allowing an image recognition means to recognize movement of an object and allowing a voice recognition means to recognize a desired direction along which a camera optical axis is desired to be moved.

SOLUTION: An information processing unit (CPU) 110 transfers sound and image information received from a microphone 140 and a camera 130 to other information processing unit via a communication line and the sound and image information transferred from other information processing unit is outputted to a speaker and a monitor 120. The CPU 110 stores an image of an object in front of a camera and when the object is moved, the image recognition means calculates the moving amount to control a moving stage 150 so as to track the

object in a way that the object is caught in the center of the screen. When the tracking is failed because the object is moved too fast or the size is too large or the like, the object makes a command in voice such as 'right' or 'left' and the CPU 110 allows the voice recognition means to recognize the command and to turn the moving stage 150 in a designated direction.

COPYRIGHT: (C)1997,JPO

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-83856

(43)公開日 平成9年(1997)3月28日

(51)Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 04 N 5/232			H 04 N 5/232	C
G 06 T 1/00			G 08 B 21/00	E
G 08 B 21/00			25/00	5 1 0 M
25/00	5 1 0		G 10 L 3/00	5 5 1 G
G 10 L 3/00	5 5 1		H 04 N 7/15	

審査請求 未請求 請求項の数1 OL (全5頁) 最終頁に続く

(21)出願番号 特願平7-229910

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(22)出願日 平成7年(1995)9月7日

(72)発明者 青柳 滋己

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72)発明者 萩田 紀博

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72)発明者 尾内 理紀夫

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(74)代理人 弁理士 秋田 収喜

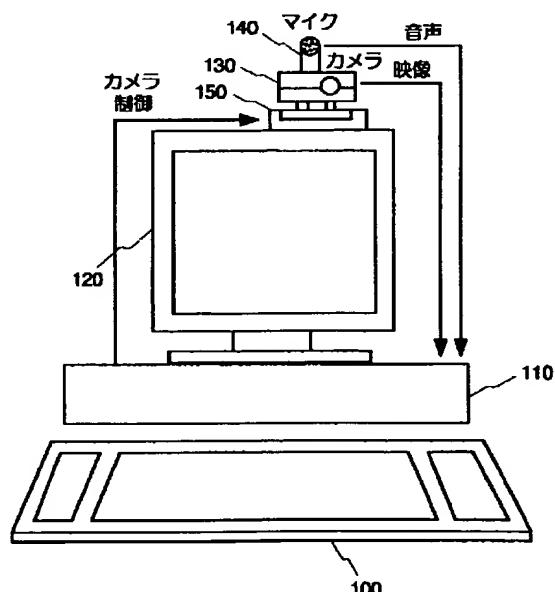
(54)【発明の名称】 インテリジェントカメラ装置

(57)【要約】

【課題】 出席者あるいは専任のオペレータ等がカメラの構図を調整することなく、カメラ自身が被写体の動きを検出し、自動的に被写体を撮影することが可能な技術を提供すること。

【解決手段】 周囲の音を集音するマイクと、画像の撮像方向あるいは撮像倍率を制御できるカメラと、前記マイクの集音した音および前記カメラが撮像した画像を入力とする情報処理装置とを有するインテリジェントカメラ装置において、前記マイクによって集音された音を認識する音認識手段と、前記カメラによって撮像された画像を認識する画像認識手段と、前記音認識手段あるいは画像認識手段のどちらか一方、または両方の出力に基づいて、前記カメラの撮像方向あるいは撮像倍率を制御し、予め定められた被写体を追従するカメラ制御手段とを備備する。

図1



1

## 【特許請求の範囲】

【請求項1】周囲の音を集音するマイクと、画像の撮像方向あるいは撮像倍率を制御できるカメラと、前記マイクの集音した音および前記カメラが撮像した画像を入力とする情報処理装置とを有するインテリジェントカメラ装置において、前記マイクによって集音された音を認識する音認識手段と、前記カメラによって撮像された画像を認識する画像認識手段と、前記音認識手段あるいは画像認識手段のどちらか一方、または両方の出力に基づいて、前記カメラの撮像方向あるいは撮像倍率を制御し、予め定められた被写体を追従するカメラ制御手段とを具備することを特徴とするインテリジェントカメラ装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、マイクと情報処理装置とを備えるインテリジェントカメラ装置に関し、特に、予め設定した被写体（物もしくは人）を自動的に追従するインテリジェントカメラ装置に適用して有効な技術に関するものである。

## 【0002】

【従来の技術】従来、テレビ会議あるいはセキュリティシステム等ではカメラ（特に、テレビカメラ）は広く使用されており、たとえば、セキュリティシステムでは管理者が遠隔操作によってカメラの視野方向を任意に変えることにより、広範囲の場所の監視あるいは特定の物体を追跡し監視するカメラ装置等が使用されている。

【0003】また、テレビ会議システムとしては、「テレミート100/50」、佐藤他、National Technical Report Vol. 40 No. 6, Dec., 1994等、そのシステムのカメラ装置の一例が紹介されている。

【0004】一方、単にカメラの撮像画像を解析したりするだけでなく、解析結果を基にカメラの視野方向や設置場所等を変更し、新たに得られた画像から更なる解析を行うという研究もなされている。

【0005】この研究の一例として、「アクティブビジョン-歴史と展望-」、久野義徳、人工知能学会誌Vol. 1. 10 No. 4, 1995に記載されるように、人工知能の分野で研究されているアクティブビジョンの分野では、撮影対象等に応じてカメラを動かし、その結果撮影された画像情報をもとにロボット等の動作を行うなどの研究がされている。

【0006】また、音声の認識等の研究も古くから行われており、たとえば、「音声研究の現状と課題」、柏谷他、日本音響学会誌Vol. 151 No. 8, 1995等に示されており、笑い声に反応してシャッターが自動的にきられるカメラが過去に製品化されている。

## 【0007】

【発明が解決しようとする課題】本発明者は、前記従来

2

技術を検討した結果、以下の問題点を見いたした。

【0008】前述する従来のテレビ会議等に用いるカメラ装置では、カメラの撮像方向や撮像倍率等のカメラの構図を調整するために、会議の出席者が自身がカメラを調整する必要があった。

【0009】このため、出席者が座る位置等を変更した場合であっても、その都度、出席者が自身がカメラの撮像方向や撮像倍率等を調整する必要があり、会議に集中できないという問題があった。

10 【0010】また、カメラが情報処理装置等と接続され、この情報処理装置に指示を行うことにより、カメラの構図を調整できるカメラ装置を使用した場合であっても、情報処理装置の操作は出席者あるいは専任のオペレータが行う必要があるという問題があった。

【0011】本発明の目的は、出席者あるいは専任のオペレータ等がカメラの構図を調整することなく、カメラ自身が被写体の動きを検出し、自動的に被写体を撮影することが可能な技術を提供することにある。

20 【0012】本発明の前記ならびにその他の目的と新規な特徴は、本明細書の記述及び添付図面によって明らかになるであろう。

## 【0013】

【課題を解決するための手段】本願において開示される発明のうち、代表的なものの概要を簡単に説明すれば、下記のとおりである。

【0014】(1) 周囲の音を集音するマイクと、画像の撮像方向あるいは撮像倍率を制御できるカメラと、前記マイクの集音した音および前記カメラが撮像した画像を入力とする情報処理装置とを有するインテリジェントカメラ装置において、前記マイクによって集音された音を認識する音認識手段と、前記カメラによって撮像された画像を認識する画像認識手段と、前記音認識手段あるいは画像認識手段のどちらか一方、または両方の出力に基づいて、前記カメラの撮像方向あるいは撮像倍率を制御し、予め定められた被写体を追従するカメラ制御手段とを具備する。

30 【0015】前述した手段によれば、例えば、画像認識手段の出力に基づいて予め定められた物体をカメラで追従する場合、まず、カメラで追従させる被写体となる物体あるいは人物の輪郭線等の概略を撮像し、情報処理装置に記憶させる。

【0016】次に、画像認識手段は情報処理装置に記憶された画像と次のフレームで撮像された画像とを比較することにより、被写体の移動量を計算し、カメラ制御手段に移動量を出力する。

【0017】カメラ制御手段は被写体の移動量に基づいて、カメラに撮影される被写体の位置（被写体の構図）が、被写体の移動前の位置となるカメラの移動量あるいは倍率を計算し、この計算値に基づきカメラを移動あるいはカメラの倍率を変化させる。

【0018】カメラの移動あるいはカメラの倍率の変化が終了したならば、再び、画像認識手段は画像を読み込み、この画像と移動前の画像にカメラの移動分あるいは撮影倍率の変化分を修正した画像とを比較し、被写体に移動がないかを調べ、移動が検出されたならば、前述する手順にしたがって再びカメラの移動あるいは撮影倍率の変化を行う。

【0019】以上に示す手順により、被写体の移動量を検出し、カメラを検出した移動量に相当する移動量だけ移動させることにより、カメラを被写体の動きに追従させる。

【0020】一方、被写体の移動量が速く、かつ、大きいために、カメラが被写体の動きに追従できない場合には、たとえば、被写体が発する「右」あるいは「左」という音声指示を音声認識手段が認識すると、音声認識手段は音声指示にしたがった方向への移動を、カメラ制御手段に出力する。

【0021】カメラ制御手段では、音声認識手段の出力に基づき、カメラの方向を予め定められた移動量だけ移動させることにより、カメラの撮影範囲内に被写体をとらえることができるので、再び、画像認識手段で被写体の移動を追従することができる。

【0022】このように、カメラに撮像される被写体の移動量を検出する画像認識手段の出力に基づいて、カメラ制御手段がカメラの撮像方向を被写体の移動量に相当する量だけ移動させることにより、カメラは被写体の動きに追従して被写体を撮像することができるので、被写体となる人物（出席者）あるいは専任のオペレータ等がカメラの撮像方向および撮像倍率を調整することなく、自動的に被写体を撮像することができる。

【0023】また、マイクに集音される音の中から、「右」、「左」という音声を認識する音声認識手段の出力に基づいて、カメラ制御手段がカメラの撮像方向を予め定められた量だけ移動させることにより、被写体が速くかつ大きく動いた場合でも被写体の方向にカメラを向けさせ、再び、被写体をとらえることができるので、声による簡単な指示だけで、被写体となる人物（出席者）あるいは専任のオペレータ等がカメラの撮像方向および撮像倍率を調整することなく、自動的に被写体を撮像することができる。

【0024】

【発明の実施の形態】以下、本発明について、発明の実施形態とともに図面を参照して詳細に説明する。

【0025】なお、発明の実施形態を説明するための全図において、同一機能を有するものは同一符号を付け、その繰り返しの説明は省略する。

【0026】図1は本発明の一実施形態（実施例）のインテリジェントカメラ装置の概略構成を示すブロック図であり、このインテリジェントカメラ装置はテレビ会議システムに用いるものである。

【0027】図1において、100はキーボード、110は情報処理装置、120はモニタ、130はテレビカメラ、140はマイク、150はカメラ取り付け台であり、キーボード100は周知のキーボードであり、情報処理装置110に文字情報を入力するための入力装置である。

【0028】情報処理装置110は、マイク140およびカメラ150から取り込まれた音および画像を、図示しないA/D変換器で「0 (L o w)」、「1 (H i g h)」のデジタル信号に変換し、この変換した音情報をおよび画像情報を後述する処理手順に従って処理した後、図示しない通信回線によって接続される図示しない他の情報処理装置に音情報および画像情報を転送する。

【0029】また、情報処理装置110は図示しない他の情報処理装置から転送されてくる音情報をD/A変換器でアナログ信号に変換し、図示しないスピーカから転送された音を出力すると共に、モニタ120から画像を出力する。

【0030】モニタ120は周知のモニタであり、情報処理装置110の映像出力信号を映像として表示する。

【0031】テレビカメラ130は、カメラ取り付け台150の移動台に取り付けられており、移動台を回転あるいは上下方向に調整することにより、テレビカメラ130の撮影方向を任意の方向に設定することができる。

【0032】また、テレビカメラ130は図示しない周知の視野角調整装置を有しており、情報処理装置110の指示により視野角調整装置は画像の撮影倍率を任意に設定することができる。

【0033】マイク140は、テレビカメラ130に固定されているので、テレビカメラ130と共に、マイク140の指向性方向を移動することができる。

【0034】カメラ取り付け台150はテレビカメラ130の一部であり、情報処理装置110の出力に基づき、上部の移動台に取り付けられたテレビカメラ130を回転あるいは上下方向に移動させることにより、テレビカメラ130の撮影方向を任意の方向に設定する。

【0035】図2は情報処理装置が有する被写体の追従制御を行うための制御部の概略構成を示すブロック図であり、210は音情報記憶回路、220は画像情報記憶回路、230は認識部、240は音認識手段、250は画像認識手段、260はカメラ制御手段、270は情報処理装置制御手段を示す。

【0036】図2において、音情報記憶回路210は、マイク140で集音された音のデジタルデータを記憶するための記憶回路であり、情報処理装置110の記憶領域の一部が割り当てられている。

【0037】画像情報記憶回路220は、カメラ130で撮像された画像のデジタルデータを記憶するための記憶回路であり、情報処理装置110の記憶領域の一部が割り当てられている。

【0038】認識手段230は、音認識手段240と画像認識手段250とを有し、音認識手段240は公知の技術を用いており、音情報記憶回路210に格納される音のデータを解析し、「右」あるいは「左」方向への移動指示がなされたか否かを後述する手順によって判断し、その結果をカメラ制御手段260に出力する。

【0039】なお、詳細については、「竹林、金澤、永田、瀬戸：「不特定話者音声自由対話システムTOSB URG「一マルチモーダル応答と音声応答キャンセルの利用」」，信学会論文誌A Vol. J77-A N o. 2, pp. 241-pp. 250, 1994」を参照されたい。

【0040】画像認識手段250は、公知の画像認識技術を用いており、画像情報記憶回路220に格納される1フレーム前と現行フレームの画像のデータを後述する手順によって解析し、その結果から、予め定められた被写体の移動量を計算し、算出された移動量をカメラ制御手段260に出力する。

【0041】なお、詳細については、「石黒：「注視に基づくロボットの視覚」，人工知能学会誌，Vol. 1 0, No. 4, pp. 10-16, July 1995」および「日本工業技術センター編：「コンピュータ画像処理入門」，pp. 118-161，総研出版発行」を参照されたい。

【0042】カメラ制御手段260は、公知の技術を用いており、画像認識手段250から出力される被写体の移動量に基づき、被写体の移動を相殺するようにカメラ130を左右方向あるいは上下方向に移動させるための信号（制御コマンド）をRS-232Cを介してカメラ130に出力し、カメラ取り付け台150の移動ステージを回転あるいは上下方向に動作させることにより、カメラ130の撮像方向を移動させる。

【0043】あるいは、カメラ130の図示しない撮影倍率調整装置を制御し、カメラ130の視野角を変化させる。

【0044】また、カメラ制御手段260は、音認識手段240から出力される右移動あるいは左移動指示に基づき、カメラ130を左右方向に予め定められた量だけ移動させるための信号をカメラ130に出力し、カメラ取り付け台150の移動ステージを回転動作させることにより、カメラ130の撮像方向を移動させる。

【0045】なお、詳細については、「佐藤、太田、北島、丸山、尾形：「新世代テレビ会議システムテレミート100/50」，National Technical Report, Vol. 40 No. 6, Dec. 1994」を参照されたい。

【0046】情報処理装置制御手段（計算機システム制御部）270は公知の技術を用いており、たとえば、「佐藤、太田、北島、丸山、尾形：「新世代テレビ会議システムテレミート100/50」，National

Technical Report, Vol. 40 No. 6, Dec. 1994」に記載されており、被写体となる出席者がキーボード100あるいは図示しないマウス等のポインティングデバイスにより行う指示でカメラ130の制御を行う。

【0047】次に、図1および図2に基づき、本実施例のインテリジェントカメラ装置の動作を、たとえば、カメラ103の前に座る被写体（ユーザー、出席者）が座る位置を図1の紙面に向かって左側に移動した場合について説明する。

【0048】まず、カメラ130が移動した被写体の画像（移動後画像）を撮像すると、この画像は情報処理装置110の図示しないA/D変換器でデジタル信号に変換された後、画像情報記憶回路220に送られ、ここで画像情報（移動後画像情報）として一時記憶されると共に、情報処理装置制御手段270から図示しない通信回線を介して接続される情報処理装置に転送される。

【0049】一方、画像認識手段250は画像情報記憶回路220から移動後画像情報を読み出し、この移動後画像情報と1フレーム前に撮像した被写体が移動する前の画像（移動前画像）の画像情報（移動前画像情報）とを後述する手順で比較する。

【0050】このとき、画像認識手段250が被写体の移動を検出すると、画像認識手段250は後述する手順に従い、カメラ130の撮影範囲内（構図内）での被写体の移動量を算出し、この移動量をカメラ制御手段260に出力する。

【0051】カメラ制御手段260は、画像認識手段250が算出した移動量に基づき、カメラ取り付け台150の移動ステージを図1の紙面に向かって左方向に回転させるための回転量を計算し、移動ステージを回転させる。

【0052】この結果、カメラ130が撮像する被写体の構図内の位置は、被写体が移動する前と同じ位置となり、カメラ130の撮影方向を被写体である出席者が座る位置を移動する度に、出席者自身がカメラ130の撮影方向を調整し直す必要がない。

【0053】以上説明したように、本実施形態によれば、カメラ130に撮像される被写体の移動量を検出する画像認識手段250の出力に基づいて、カメラ制御手段260がカメラ130の撮影方向を被写体の移動量に相当する量だけ移動させることにより、カメラ130は被写体の動きに追従して被写体を撮像することができる。出席者あるいは専任のオペレータ等がカメラ130の撮影方向あるいは撮像倍率を調整することなく、自動的に被写体を追従し撮像することができる。

【0054】また、マイク140に集音される音の中から、「右」、「左」という音声を認識する音認識手段240の出力に基づいて、カメラ制御手段260がカメラ130の撮像方向を予め定められた量だけ移動させるこ

とにより、被写体が速くかつ大きく動いた場合でも被写体の方向にカメラ130を向けさせ、再び、被写体をとらえることができるので、声による簡単な指示だけで、出席者あるいは専任のオペレータ等がカメラ130の撮像方向あるいは撮像倍率を調整することなく、自動的に被写体を撮像することができる。

【0055】なお、本実施形態においては、音認識手段240によって被写体である出席者の発する「右」あるいは「左」という音声を認識し、その結果に基づいてカメラ制御手段260がカメラ130を回転させる構成であるが、音認識手段240とは別に、たとえば、マイク140を2本用いて、周囲の音をステレオで集音し、この音の位相の差から音源方向を特定する音源方向特定手段を設け、その出力に基づきカメラ130の撮像方向を制御することにより、音がした方にカメラ130を向かせることができる。

【0056】また、前述するような音源方向特定手段と、高速に撮像方向を移動できるカメラ130とを有するインテリジェントカメラ装置をセキュリティーシステムに用いることにより、音がした方の画像をより明瞭に撮像することができる。

【0057】以上、本発明者によってなされた発明を、前記発明の実施形態に基づき具体的に説明したが、本発

明は、前記発明の実施形態に限定されるものではなく、その要旨を逸脱しない範囲において種々変更可能であることは勿論である。

【0058】

【発明の効果】本願において開示される発明のうち代表的なものによって得られる効果を簡単に説明すれば、下記の通りである。

【0059】カメラ自身が被写体の動きを検出し、自動的に被写体を追従して撮影することができるので、出席者あるいは専任のオペレータ等がカメラの構図を調整する必要がなくなる。

【図面の簡単な説明】

【図1】本発明の一実施形態であるインテリジェントカメラ装置の概略構成を示すブロック図である。

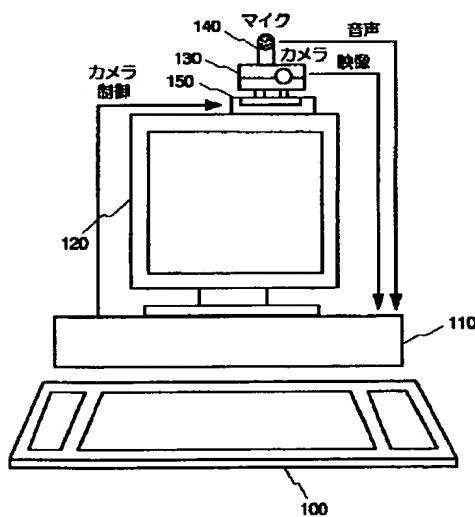
【図2】情報処理装置が有する被写体の追従制御を行うための制御部の概略構成を示すブロック図である。

【符号の説明】

100…キーボード、110…情報処理装置、120…モニタ、130…テレビカメラ、140…マイク、150…カメラ取り付け台、210…音情報記憶回路、220…画像情報記憶回路、230…認識部、240…音認識手段、250…画像認識手段、260…カメラ制御手段、270…情報処理装置制御手段。

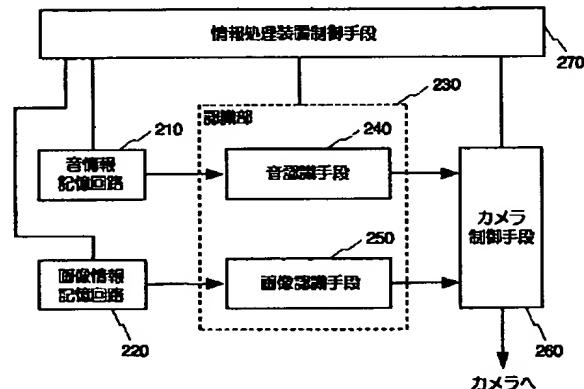
【図1】

図1



【図2】

図2



フロントページの続き

(51) Int. Cl. 6

H04N 7/15  
7/18

識別記号

府内整理番号

F I

H04N 7/18  
G06F 15/64

技術表示箇所

E  
325B